

PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (<http://bmjopen.bmj.com/site/about/resources/checklist.pdf>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Association between neighbourhood walkability and metabolic risk factors influenced by physical activity: a cross-sectional study of adults in Toronto, Canada
AUTHORS	Loo, C K Jennifer; Greiver, Michelle; Aliarzadeh, Babak; Lewis, Daniel

VERSION 1 - REVIEW

REVIEWER	Mark Green University of Liverpool, UK
REVIEW RETURNED	02-Sep-2016

GENERAL COMMENTS	<p>Thank you for the opportunity to review the paper. The paper explores the association between neighbourhood walkability and objective measures of metabolic risk factors. The paper is mostly fine and my comments are mainly minor. Largely it needs a bit more detail and clarity in places.</p> <p>Main comments:</p> <ul style="list-style-type: none">- The introduction is a little short - concise is always good - but it could do with a little more detail in places. Some readers of the journal may be less familiar with why you would want to look at neighbourhoods, both conceptually and for policy reasons so adding in a little more detail here would help- The methodology needs more detail on why you selected your variables. I don't have any qualms myself with their selection but they still need some justification. This is particularly important for the covariates e.g. why include ethnic concentration? Please add for all variables.- I was a little unsure why you would split up the walkability measure into quartiles - by transforming it from a continuous variable you are losing information and therefore not describing your data precisely.- You say that you do not use multi-level models because "There were insufficient observations within each neighbourhood" (page 6, lines 6-8). You have a sample of 78 023 cases and note that there were 140 neighbourhoods. 78 023 divided by 140 is 557.3 - so on average you should have 557.3 cases per area, which should be enough to estimate a multi-level model. Of course, there may be little variability between areas, which might be an interesting finding yourself, but I do not buy your current justification.- You disaggregate your results by age group which I liked - how about by sex?- The key results section is brief. Similarly to the content in the abstract, saying that there was an association between walkability and blood pressure as a main finding is a little misleading as it was only consistent for one age group.- Please add more references to support your comments in the
-------------------------	---

	<p>limitations</p> <ul style="list-style-type: none"> - The association between BMI and walkability was stronger at younger ages - how so? - I was interested in why only middle aged individuals demonstrated an association between walkability and the two blood pressure measures but you gloss over this in the discussion. Can we add some explanation please? - If you are going to discuss 'equity' (page 14, lines 9-13) then might be worth talking about social inequalities too. Suggest that you either expand or drop the paragraph. <p>Nit-picking comments:</p> <ul style="list-style-type: none"> - The keywords could be better - The conclusions in the abstract (p2, lines 44-46) are dull and repetitive of the results. Say something new! - Please avoid starting paragraphs with 'this' as is vague - Page 7 line 11-12 - you say the range of walkability values was 42-100 but on the map it only extends to 99. - Table 2 contains a significance test but this is not mentioned in the table or the methodology. Please add.
--	--

REVIEWER	F Mueller-Riemenschneider National University of Singapore
REVIEW RETURNED	19-Sep-2016

GENERAL COMMENTS	<p>Manuscript: bmjopen-2016-013889 Title: ASSOCIATION BETWEEN NEIGHBOURHOOD WALKABILITY AND METABOLIC RISK FACTORS INFLUENCED BY PHYSICAL ACTIVITY: A CROSS-SECTIONAL STUDY OF ADULTS IN TORONTO, CANADA</p> <p>General comments This study investigates the association between neighbourhood walkability and metabolic risk factors in a large sample of primary care patients in Toronto. The study addresses a topic of public health relevance and expands on previous research in the field. The manuscript is generally well written. The analysis and study has some methodological issues, however, that have implications for the conclusions drawn.</p> <p>Specific comments Abstract</p> <ul style="list-style-type: none"> - The authors should provide details on statistical analysis performed - Results: provide information about the study sample - What is the rationale for comparing only highest vs. lowest quartile? - Use of abbreviations in abstract may allow authors to provide more details, listing all the metabolic outcomes multiple times takes up a lot of space. On the other hand, BMI has not been explained - it is not clear whether results are adjusted for possible confounders - Given the sample size it is surprising that no stratified analysis is presented/performed - Why are the conclusion only highlighting associations with BMI and BP, but not other outcomes? Where they considered not to be clinically meaningful? <p>Introduction</p> <ul style="list-style-type: none"> - The introduction is very brief. It would be helpful if the authors could provide, for instance, further information on the validity of the walkscore (has it ever been compared to other established
-------------------------	---

	<p>walkability indices?). In addition, the authors should clearly describe existing studies that have explored this topic and potential gaps in the existing evidence</p> <p>Methods</p> <ul style="list-style-type: none"> - It would be helpful if the authors could provide further information about the Canadian primary care setting. How likely is the enrolment from such records going to result in selection bias or in the inclusion of overly sick individuals that are not representative of the general population? - Covariates: does the EMR have no further individual level information, for instance on ethnicity, education, work status? How about other health behaviours, such as physical activity, alcohol intake? (diet is mentioned to be missing by the authors). The lack of including this information could be a major source for residual confounding. - The authors should justify their choice to only compare Q1 vs. Q4, why not including the others and investigate dose response relationships? - The authors should also justify a stratification by age, but not by other potentially relevant factors (e.g. gender) <p>Results</p> <ul style="list-style-type: none"> - Table 1: are further individual level factors available (behavioural, SES etc.)? why do the authors not provide information about differences between compared groups (p-values) - Table 2: again, it would be interesting to see information from all quartiles. Please provide explanations of abbreviations underneath - The authors may want to reduce word count by reducing overlap between information in text and tables - Given the sample size I feel the analysis could be strengthened by providing further analysis (e.g. further stratification, use of walkscore as a continuous measure) <p>Discussion</p> <ul style="list-style-type: none"> - Page 12, line 17: the authors acknowledge the limitation in generalising primary care patients. It would be interesting to understand to what extent the population is similar to the general adult population in the region. - The authors acknowledge two major limitations of the study, cross-sectional study design and the fact that they have not controlled for a large number of individual level factors. Neighbourhood level factors (especially considering the scale of the neighbourhoods described here) can probably not fully capture such differences which can have major implications for the findings and conclusions. Under these circumstances a much more careful interpretation seems warranted. An additional limitation is the chosen definition of neighbourhood walkability (e.g. scale of neighbourhood area considered, type of index used), which should be critically discussed and compared with other studies in the field in more detail
--	---

REVIEWER	J.Lakerveld VU University medical center Amsterdam
REVIEW RETURNED	23-Dec-2016

GENERAL COMMENTS	<p>The manuscript reports on an interesting analysis linking patient record data with walkscore data.</p> <p>Although in general I think the manuscript includes most of the necessary information there are some considerations and issues that should be addressed, or addressed better. Including:</p>
-------------------------	---

	<ul style="list-style-type: none"> - The potential time gap between the assessment of outcomes and the derived walkscore data; - the Walk Score being a composite measure. The complexities of composite measures may reduce the transferability for public health and planning practitioners and policymakers; - the lack of individual-level SES indicators such as education or income; - a description of the external validity and potential selection bias, as a specific patient population is used; - the limitations of using administrative neighbourhood boundaries, and not including other relevant environments of participants (such as work environment) <p>Furthermore it is unclear to me why adults with less than two visits were excluded.</p>
--	--

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Reviewer Name

Mark Green

Institution and Country

University of Liverpool, UK

Please state any competing interests or state 'None declared':

None declared

Please leave your comments for the authors below Thank you for the opportunity to review the paper. The paper explores the association between neighbourhood walkability and objective measures of metabolic risk factors. The paper is mostly fine and my comments are mainly minor. Largely it needs a bit more detail and clarity in places.

Main comments:

- The introduction is a little short - concise is always good - but it could do with a little more detail in places. Some readers of the journal may be less familiar with why you would want to look at neighbourhoods, both conceptually and for policy reasons so adding in a little more detail here would help

***Response:

Further details have been added to the introduction to expand on the concept of utilitarian walking and to draw the link between neighbourhood design and health.

- The methodology needs more detail on why you selected your variables. I don't have any qualms myself with their selection but they still need some justification. This is particularly important for the covariates e.g. why include ethnic concentration? Please add for all variables.

***Response:

We have provided additional explanations for the selection of outcome variables and covariates in the Methods section. The Ontario Marginalization Index is a deprivation-based index that includes four dimensions of marginalization, of which ethnic concentration is one. Ethnic concentration is relevant to area-level marginalization, and, as with all the domains of the index, it has impacts on health outcomes. Further information on the development, validation, and application of this index can be found in the related citations provided in the manuscript.

- I was a little unsure why you would split up the walkability measure into quartiles - by transforming it from a continuous variable you are losing information and therefore not describing your data precisely.

***Response:

Although the Walk Score is a continuous variable, it is not strictly an interval variable – a score of 40 is not, in a sense, twice as good as a score of 20, and we cannot be sure that an increase in Walk Score from 21 to 22 is the same as an increase from 80 to 81. As such, we prefer to group the data into quartiles, so that we can compare the relative difference in effects between residents of the most walkable neighbourhoods and residents of the least walkable neighbourhoods.

Broadly, our use of quantiles follows the approach of other social statistics aiming to give a relative overview of neighbourhood circumstances. An example would be the Index of Multiple Deprivation (IMD), where it is similarly appropriate to present quantiles, but where reporting the score is potentially misleading.

- You say that you do not use multi-level models because "There were insufficient observations within each neighbourhood" (page 6, lines 6-8). You have a sample of 78 023 cases and note that there were 140 neighbourhoods. 78 023 divided by 140 is 557.3 - so on average you should have 557.3 cases per area, which should be enough to estimate a multi-level model. Of course, there may be little variability between areas, which might be an interesting finding yourself, but I do not buy your current justification.

***Response:

We did consider the use of multi-level models and our decision to proceed with a non-hierarchical approach was based on two factors. Firstly, for most of the health outcome variables, we found that the minimum number of observations in a neighbourhood was low – 10-12 observations, and as low as 6 observations in a neighbourhood for one outcome variable. Secondly, as we stated in the Statistical Analyses section, upon calculating the intra-class correlation coefficients for each outcome variable, we noted that very little of the total variance was accounted for by clustering within neighbourhoods. As such, we felt that a non-hierarchical approach was valid and reasonable.

- You disaggregate your results by age group which I liked - how about by sex?

***Response:

Age stratification was an a priori decision in our study, and we outlined the important reasons for this in the Methods section under Statistical Analysis. Sex stratification would represent a post hoc analysis. While we fully agree that this would be an interesting exploratory analysis, we feel that, with nine outcome variables already stratified by age, the further addition of sex stratification would detract from the clarity of the main results.

- The key results section is brief. Similarly to the content in the abstract, saying that there was an association between walkability and blood pressure as a main finding is a little misleading as it was only consistent for one age group.

***Response:

We have modified the key results and the abstract conclusion to improve clarity and to highlight that it is the clinical significance of the results that we wish to emphasize (mean differences in BMI for all age groups, and mean differences in blood pressure for middle-aged adults).

- Please add more references to support your comments in the limitations

***Response:

We have expanded this section to include additional references to other studies.

- The association between BMI and walkability was stronger at younger ages - how so?

***Response:

Our previous statement was intended to highlight the greater magnitude of the difference in mean BMI in younger age groups, not the strength of the association. We have modified this paragraph to improve clarity (see response above).

- I was interested in why only middle aged individuals demonstrated an association between walkability and the two blood pressure measures but you gloss over this in the discussion. Can we add some explanation please?

***Response:

We have provided some possible explanations as to why the association between walkability and blood pressure was observed in the middle age group and not others. This is accompanied by a brief description of the epidemiology of hypertension in Canada which provides the necessary context.

- If you are going to discuss 'equity' (page 14, lines 9-13) then might be worth talking about social inequalities too. Suggest that you either expand or drop the paragraph.

***Response:

We agree that it is beyond the scope of this paper to discuss social inequalities at length, and there is excellent literature on social inequalities elsewhere. Our main point here is that the built environment is a determinant of health, and one which may vary considerably, even within a single city such as Toronto. As such, public health practitioners and policy makers who seek to address the determinants of health and promote health equity should remember that unequal access to built environment interventions (e.g. public transportation, safe pedestrian infrastructure) may contribute to unequal health impacts.

Nit-picking comments:

- The keywords could be better

***Response:

We have added some additional keywords, including "active transportation", "metabolic risk" and "utilitarian walking".

- The conclusions in the abstract (p2, lines 44-46) are dull and repetitive of the results. Say something new!

***Response:

We have modified our conclusion slightly based on comments from all reviewers. The main thrust of our conclusion is to emphasize that the magnitude of differences observed for BMI are not just statistically significant, but also clinically significant; furthermore, these clinically significant differences in mean BMI were observed across all adult age groups.

- Please avoid starting paragraphs with 'this' as is vague

***Response:

We have changed the structure of the sentence in question to improve clarity.

- Page 7 line 11-12 - you say the range of walkability values was 42-100 but on the map it only extends to 99.

***Response:

We have made the correction. The Walk Score range extends to 100 but, as you pointed out, the range of Walk Scores for Toronto neighbourhoods extends only to 99.

- Table 2 contains a significance test but this is not mentioned in the table or the methodology. Please add.

***Response:

We have clarified in both the methods and in the table that t-tests were used.

Reviewer: 2

Reviewer Name

F Mueller-Riemenschneider

Institution and Country

National University of Singapore

Please state any competing interests or state 'None declared':

none declared

Please leave your comments for the authors below

Manuscript: bmjopen-2016-013889

Title: ASSOCIATION BETWEEN NEIGHBOURHOOD WALKABILITY AND METABOLIC RISK FACTORS INFLUENCED BY PHYSICAL ACTIVITY: A CROSS-SECTIONAL STUDY OF ADULTS IN TORONTO, CANADA

General comments

This study investigates the association between neighbourhood walkability and metabolic risk factors in a large sample of primary care patients in Toronto. The study addresses a topic of public health relevance and expands on previous research in the field. The manuscript is generally well written. The analysis and study has some methodological issues, however, that have implications for the conclusions drawn.

Specific comments

Abstract

- The authors should provide details on statistical analysis performed
- Results: provide information about the study sample
- What is the rationale for comparing only highest vs. lowest quartile?
- Use of abbreviations in abstract may allow authors to provide more details, listing all the metabolic outcomes multiple times takes up a lot of space. On the other hand, BMI has not been explained
- it is not clear whether results are adjusted for possible confounders
- Given the sample size it is surprising that no stratified analysis is presented/performed
- Why are the conclusion only highlighting associations with BMI and BP, but not other outcomes? Where they considered not to be clinically meaningful?

***Response:

We have clarified the abstract to specify that multivariable linear regression models were used, and have listed the covariates. We have also highlighted that age-stratification was performed and presented these results for BMI. We have clarified that BMI stands for body mass index. Information about the study sample is presented in the "Participants" section of the abstract.

Whilst we compared all quartiles of walkability (see Supplementary Table), we reported results from the highest and lowest quartiles because we were interested in establishing whether there was a plausible clinical difference in metabolic risk factors based on the range of neighbourhood walkability experienced by Toronto residents. As such we establish that those in the least walkable neighbourhoods, controlling for other individual factors and other characteristics of neighbourhood circumstance, do differ in metabolic risk factors compared to those in the most walkable areas.

We have modified our conclusion slightly to emphasize the BMI results, which are clinically meaningful. While statistically significant differences were observed for other health measures, the magnitude of these differences were small, and would not be considered to be very meaningful in the clinical setting. On the other hand, the magnitude of the differences in mean BMI are clinically significant across all age groups, and we have therefore highlighted the BMI results in our conclusion.

Introduction

- The introduction is very brief. It would be helpful if the authors could provide, for instance, further information on the validity of the walkscore (has it ever been compared to other established walkability indices?). In addition, the authors should clearly describe existing studies that have explored this topic and potential gaps in the existing evidence

***Response:

In order to balance the need for a concise introduction with detailed descriptions of existing gaps in the evidence, we have highlighted broad findings from the current literature in the Introduction, and we have articulated the main gap that this study addresses (i.e. the use of objectively measured metabolic risk factors as outcome measures).

A more detailed description of the Walk Score and its validity as a walkability measure is provided in the Methods section. We have also provided additional examples of its association to health outcomes in the Canadian context.

Given the suite of outcome measures examined in our study, we agree that findings from existing studies should be clearly described. Under the section, Findings in Relation to Other Studies, we have structured our discussion to comment on our results in the context of other studies for each group of outcomes.

Methods

- It would be helpful if the authors could provide further information about the Canadian primary care setting. How likely is the enrolment from such records going to result in selection bias or in the inclusion of overly sick individuals that are not representative of the general population?

***Response:

As requested, we have provided additional information about the Canadian primary care setting in the Methods section. Because the study used primary care patients, we also comment on the implications for generalizability in our Discussion, under Strengths and Limitations.

- Covariates: does the EMR have no further individual level information, for instance on ethnicity,

education, work status? How about other health behaviours, such as physical activity, alcohol intake? (diet is mentioned to be missing by the authors). The lack of including this information could be a major source for residual confounding.

***Response:

With respect to covariates, although EMR data is an excellent source of objectively measured health outcome variables, EMR data on ethnicity, education, work status, and certain health behaviours are unstructured, missing for most patients, and of low quality. We therefore used residential postal codes to generate neighbourhood-level indicators. Neighbourhood-level covariates are described in the Methods section. We have expanded the Strengths and Limitations discussion to comment in greater detail on the implications for residual confounding, particularly for physical activity.

Broadly, because our data is derived from EMRs in the primary care context, it is less rich in individual characteristics than a specifically designed sample survey, which we accept as an inevitable weakness. Despite this, we feel that the EMR dataset has some considerable advantages. As a routinely collected administrative record of interactions by enrolled patients with primary care physicians it offers an incredibly rich insight into the population prevalence of a range of important health risks and outcomes. There are few sample surveys that offer the same quantity of structured data of the health of individuals, particularly accounting for anthropometric and biomarker measures, and information about prescriptions. Further, whilst we have used the data in a cross-sectional and largely exploratory manner here, we feel that such datasets will ultimately yield important longitudinal data. Should data linkage become an option we may, in future, be able to account for other individual-level data, such as those factors you mention.

- The authors should justify their choice to only compare Q1 vs. Q4, why not including the others and investigate dose response relationships?

***Response:

Please see our earlier response on the use of quartiles. On the subject of dose response, we feel that the state of the research is such that asserting a dose-response relationship would be premature. Certainly we observe an association, and that this relationship is graded to some extent according to the walkability of peoples' residential neighbourhoods. However, our work is exploratory, and little is currently known about the mechanisms by which walkability is either embodied or drives different health behaviours. As you have suggested, the association observed may not be clear cut, owing to the effect of residual confounding and a body of literature that is currently somewhat equivocal on the topic, albeit of variable quality. Nonetheless, we feel that our findings are useful from both a clinical and a population perspective. We add to a growing body of high-quality evidence for the positive effect of neighbourhood walkability in cardiovascular and metabolic health, which may help to motivate more detailed work on mechanism and health behaviours. Moreover, we demonstrate the utility of EMR data to the observation and exploration of health concerns at the intersection of clinical and population health relevance.

- The authors should also justify a stratification by age, but not by other potentially relevant factors (e.g. gender)

***Response:

Age stratification was an a priori decision in our study, and we outlined the reasons for this in the Methods section under Statistical Analysis. Most importantly, the clinical approach to the screening, diagnosis, and treatment of adults for each of our outcome measures does differ depending on these broad age categories (much less so for sex). As a result, we felt it was important that age-stratified results be presented for appropriate interpretation. Sex stratification would represent a post hoc analysis. While we fully agree that this would be an interesting exploratory analysis, we feel that, with nine outcome variables already stratified by age, the further addition of sex stratification would detract

from the clarity of the main results.

Results

- Table 1: are further individual level factors available (behavioural, SES etc.)? why do the authors not provide information about differences between compared groups (p-values)

***Response:

Please see our earlier response with respect to EMR data and individual level factors.

The data presented in Table 1 is descriptive only, to give the reader a sense of the characteristics of the two groups, and for the total study population. P-values are not provided because the intent of the study was not to test for differences in these characteristics; rather, information in Table 1 provides the basis for considering these factors for adjustment, and including them as covariates in our regression models.

- Table 2: again, it would be interesting to see information from all quartiles. Please provide explanations of abbreviations underneath

***Response:

Table 2 has been modified to include the full wording for each outcome variable. We agree that results from all quartiles would be important to present; due to spacing limitations, these are presented only for the fully adjusted analyses, in the Supplementary Table.

- The authors may want to reduce word count by reducing overlap between information in text and tables

***Response:

We have further condensed the text in the Results section by reducing the overlap in information also found in the tables.

- Given the sample size I feel the analysis could be strengthened by providing further analysis (e.g. further stratification, use of walkscore as a continuous measure)

***Response:

Please see our earlier response with respect to the a priori reasoning behind age stratification. While we agree that post hoc stratification by other factors may be interesting in an exploratory sense, due to space limitations, the presentation of additional data would detract from the clarity of our main results.

We chose to compare the highest to the lowest quartile of walkability because while the Walk Score is a continuous variable, it is not strictly an interval variable. A score of 40 is not, in a sense, twice as good as a score of 20, and we cannot be sure that an increase in Walk Score from 21 to 22 is the same as an increase from 80 to 81. As such, we chose to group the data into quartiles. By doing so, we assumed that there is relative homogeneity within groups compared to between groups, allowing us to compare for the relative difference in effects between those people who live in the most walkable neighbourhoods compared to the least walkable neighbourhoods.

Discussion

- Page 12, line 17: the authors acknowledge the limitation in generalising primary care patients. It would be interesting to understand to what extent the population is similar to the general adult

population in the region.

***Response:

As briefly mentioned in the paragraph you indicate, we note that the study population was older and had a greater proportion of women, compared to the general Toronto population. This is a known feature of CPCSSN patient populations and underscores the importance of sex and age adjustment when using CPCSSN data. We also expanded on other aspects of generalizability, such as the prevalence of comorbidities in the study population compared to the general population. Work is under way to explore the representativeness of CPCSSN patients with respect to other important characteristics.

- The authors acknowledge two major limitations of the study, cross-sectional study design and the fact that they have not controlled for a large number of individual level factors. Neighbourhood level factors (especially considering the scale of the neighbourhoods described here) can probably not fully capture such differences which can have major implications for the findings and conclusions. Under these circumstances a much more careful interpretation seems warranted. An additional limitation is the chosen definition of neighbourhood walkability (e.g. scale of neighbourhood area considered, type of index used), which should be critically discussed and compared with other studies in the field in more detail

***Response:

We have provided additional elements in the Strengths and Limitations section of the Discussion in response to these and other reviewer comments.

We fully agree that the scale at which neighbourhood walkability is measured is important, particularly if it is to be considered as a built environment intervention. We make reference to this concept in the discussion of the BMI results, in the section on Findings in Relation to Other Studies, and we also revisit the importance of spatial scale in our Conclusion. Further examination of the merits of different scales and indices of neighbourhood walkability would greatly contribute to the evidence base in this field, but is beyond the scope of our present study and the word limits of this paper.

Reviewer: 3

Reviewer Name

J.Lakerveld

Institution and Country

VU University medical center Amsterdam

Please state any competing interests or state 'None declared':

non declared

Please leave your comments for the authors below The manuscript reports on an interesting analysis linking patient record data with walkscore data.

Although in general I think the manuscript includes most of the necessary information there are some considerations and issues that should be addressed, or addressed better. Including:

- The potential time gap between the assessment of outcomes and the derived walkscore data;

***Response:

We have clarified that the health outcome measures were taken in the 3-year period between 2012 and 2014 and that Walk Scores current as of 2014-2015 were used. We believe that this represents an adequately narrow time frame for the purpose of our cross-sectional study. Given the nature of the

Walk Score as a composite walkability index, it is unlikely that, within a three-year period, a Toronto neighbourhood Walk Score would change so significantly as to place the neighbourhood in an opposite quartile of walkability. We believe that the risk of misclassification is low with our 3-year time frame.

- the Walk Score being a composite measure. The complexities of composite measures may reduce the transferability for public health and planning practitioners and policymakers;

***Response:

There is always a balance to be struck between more general characterisations and approaches which seek to draw out the 'active ingredients' in population health research. The Walk Score does combine a range of factors into something of a "black box", which makes the specification of interventions or programmes of behaviour change difficult. However, despite it being a composite measure, the score itself is simple, understandable, and easily obtained. Indeed, in addition to researchers, a number of public health and planning practitioners in North America are already using Walk Scores to communicate the impacts of infrastructure changes. In our Methods section we make reference to the validity of the Walk Score and have provided additional examples of its association to health outcomes in the Canadian context. We also wish to highlight that, as a simple and easily obtained indicator, the Walk Score may provide physicians with some useful environmental context when optimizing the metabolic and cardiovascular risk of patients. Encouraging a patient towards greater mobility may have different effects depending on the residential context of the patient. This is certainly something that is worth studying in a diagnostic sense, alongside work which seeks to further atomise the environment to develop our understanding of the contextual drivers of cardiovascular and metabolic disease.

- the lack of individual-level SES indicators such as education or income;

***Response:

Although EMR data is an excellent source of objectively measured health outcome variables, EMR data on individual SES indicators are unstructured, missing for most patients, and of low quality. We therefore used residential postal codes to generate neighbourhood-level indicators. Neighbourhood-level covariates are described in the Methods section.

Broadly, because our data is derived from EMRs in the primary care context, it is less rich in individual characteristics than a specifically designed sample survey, which we accept as an inevitable weakness. Despite this, we feel that the EMR dataset has some considerable advantages. As a routinely collected administrative record of interactions by enrolled patients with primary care physicians it offers an incredibly rich insight into the population prevalence of a range of important health risks and outcomes. There are few sample surveys that offer the same quantity of structured data of the health of individuals, particularly accounting for anthropometric and biomarker measures, and information about prescriptions. Further, whilst we have used the data in a cross-sectional and largely exploratory manner here, we feel that such datasets will ultimately yield important longitudinal data. Should data linkage become an option we may, in future, be able to account for other individual-level data, such as those factors you mention.

- a description of the external validity and potential selection bias, as a specific patient population is used;

***Response:

In the Strengths and Limitations section of the Discussion, we note that the study population was older and had a greater proportion of women, compared to the general Toronto population. This is a known feature of CPCSSN patient populations and underscores the importance of sex and age adjustment when using CPCSSN data. We also expand on other aspects of generalizability, such as

the prevalence of comorbidities in the study population compared to the general population. Work is under way to explore the representativeness of CPCSSN patients with respect to additional characteristics.

- the limitations of using administrative neighbourhood boundaries, and not including other relevant environments of participants (such as work environment)

***Response:

The use of administrative neighbourhood boundaries has both advantages and disadvantages. On the one hand, neighbourhood-level covariates are generally easily obtained when the geographical areas of study are administrative ones. Data presented according to administrative boundaries also tends to be more meaningful to policy makers, whose jurisdiction and levers of intervention are similarly bounded by administrative borders. However, there are also limitations. For example, the scale of administrative areas do not necessarily reflect the true spatial scale at which neighbourhood walkability may realize any health effects. We make reference to this concept in our discussion of the BMI findings and we also revisit the importance of spatial scale in our Conclusion.

We acknowledge that our study does not include work- or leisure-related physical activity. As noted previously in our responses, such information is not captured in EMR data. We have expanded our limitations section and specifically commented on leisure physical activity and implications for the study results. If levels of work-related physical activity do not vary according to neighbourhood walkability (and we have no reason to believe that they would), we would similarly expect that confounding due to work-related physical activity would bias the results toward the null and lead to underestimation of effects in the present study.

Furthermore it is unclear to me why adults with less than two visits were excluded.

***Response:

By excluding patients with less than two visits, this excludes transient patients from the study sample for whom data quality and completeness may be low. EMR data depends on clinician input of data, and this is in turn influenced by the establishment of a relationship between patient and provider.

Transient patients who have only had a single encounter with the CPCSSN physician are more likely to have incomplete records, particularly with respect to lifestyle factors such as smoking status, and aspects of the medical history such as comorbidities and medications.

VERSION 2 – REVIEW

REVIEWER	Mark Green University of Liverpool, UK
REVIEW RETURNED	06-Feb-2017

GENERAL COMMENTS	Addressed issues - cheers.
-------------------------	----------------------------

REVIEWER	Lakerveld, Jeroen VUmc The Netherlands
REVIEW RETURNED	09-Feb-2017

GENERAL COMMENTS	My previous comments were addressed adequately.
-------------------------	---